AD A075663

LEVEL ~~III~~ 12

Improved ARTEP Methods for Unit Evaluation

VOLUME VI: CONVENTIONAL ARTEP MISSIONS AND ENGAGEMENT SIMULATIONS: AN EXAMINATION OF OPTIONS

by

M. Dean Havron and Leslie H. McFarling
Human Sciences Research, Inc.

and

Robert G. Wanschura
U.S. Army Research Institute for the
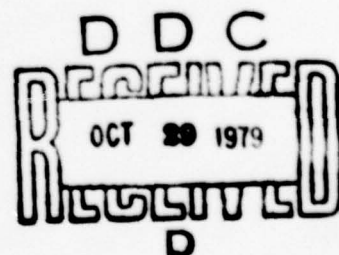Behavioral and Social Sciences

April 1979

DDC FILE COPY

Prepared for

ari

D D C
RECEIVED
OCT 29 1979
D

79 10 26 032

U. S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the
Deputy Chief of Staff for Personnel

JOSEPH ZEIDNER
Technical Director

WILLIAM L. HAUSER
Colonel, U S Army
Commander

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
TR-79-A24	ARI	

4. TITLE *(and Subtitle)*
IMPROVED ARTEP METHODS FOR UNIT EVALUATION:
VOLUME VI: CONVENTIONAL ARTEP MISSIONS AND
ENGAGEMENT SIMULATIONS: AN EXAMINATION OF OPTIONS

5. TYPE OF REPORT & PERIOD COVERED
Final Report
31 Jan 78 - 30 Apr 79

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
M. Dean Havron, Leslie H. McFarling, and
Robert G. Wanschura

8. CONTRACT OR GRANT NUMBER(s)
DAHC 19-77-C-0001

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Human Sciences Research, Inc.
7710 Old Springhouse Road

10. PROGRAM ELEMENT, PROJECT, TASK
AREA & WORK UNIT NUMBERS
2Q163743A773

11. CONTROLLING OFFICE NAME AND ADDRESS
U.S. Army Training and Doctrine Command (TRADOC)
Fort Monroe, VA 23651

12. REPORT DATE
April 1979

13. NUMBER OF PAGES
28

14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*
U.S. Army Research Institute for the Behavioral
and Social Sciences, 5001 Eisenhower Avenue
Alexandria, Virginia 22333

15. SECURITY CLASS. *(of this report)*
Unclassified

15a. DECLASSIFICATION/DOWNGRADING
SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*
--

Improved Army Training and Evaluation
Program (ARTEP) Methods for Unit Evaluation.
Volume VI. Conventional ARTEP Missions and
Engagement Simulations: An Examination of
Options.

18. SUPPLEMENTARY NOTES

Research monitored technically by Robert G. Wanschura, Engagement Simulation
Technical Area, Army Research Institute

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Army Training and Evaluation Program (ARTEP)
Engagement Simulation
REALTRAIN
MILES

Unit Performance Criteria
After-Action Review
Field Exercise Observation/
Control Systems

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Differences between the philosophies and procedures of engagement simulation
training methods and conventional ARTEP field exercises generate several issues
if both are to be integrated into exercises where unit performance measurement
and unit evaluation is desired. This report examines four issues: development
of accurate, comprehensive criteria and measures of unit tactical performance;
structure and functions of evaluator/controller teams for data collection and
performance rating; reduction/integration of ES-generated data and conventional
ARTEP Test and Evaluation Outline data; and use of collected data to establish

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

20. training objectives and assess progress in training programs. Primary audiences for this report are researchers and TRADOC personnel responsible for development of ES and ARTEP training/evaluation methodologies.

Accession For

NTIS GRA&I	
DDC TAB	
Unannounced	
Justification	

By

Distribution/

Availability Codes

Dist.	Avail and/or special
A	

TABLE OF CONTENTS

CONVENTIONAL ARTEP MISSIONS AND
ENGAGEMENT SIMULATIONS: AN EXAMINATION OF OPTIONS

1

CHAPTER 1
INTRODUCTION/BACKGROUND

1-1. INTRODUCTION

Under ARI sponsorship, HSR is completing a two-phase research study, the objectives of which is to provide guidance and means for improving ARTEP exercises for maneuver units. This volume is one of six reports resulting from the study. The purpose of this volume is to better define and explore issues involved in the integration of engagement simulation into the Army's program for unit field training.

This report discusses in order:

- Development of conventional and engagement simulation approaches to unit training.

- Steps in unit performance evaluation as conducted by conventional methods and by engagement simulation.

- Issues that need to be resolved in contemplation of inclusion of ES into ARTEP.

Since engagement simulation methods are still under development, this report is intended for the research community and Army organizations responsible for collective training methods.

1-2. BACKGROUND OF ARTEP

In 1971, the Army started to develop a family of training materials referred to as Army Training and Evaluation Programs (ARTEPs). The objective was to address needs for collective training in a peacetime environment. The ARTEP concepts may be thought of as broad umbrellas that can incorporate a variety of training concepts, methods, and techniques. These concepts and techniques take specific form in ARTEP documents which provide some five chapters of general instructions followed by Training and Evaluation Outlines (T&EOs).

3

This current two-phase study is concerned with Army guidance for Mechanized Infantry units which provides general instructions and T&EOs for maneuver elements for squads, platoons, companies and battalions. Guidance materials draw heavily from Instructional System Development and Skill Qualification Tests prepared for training individuals in MOS skills. This is reflected in the format of their Training and Evaluation Outlines (T&EOs).

Each T&EO module specifies training objectives, the conditions under which the mission is performed, and the primary training/evaluation standards upon which the element will be evaluated. (Suggested support requirements, i.e., evaluator personnel, ammunition, aggressors, maneuver area, etc., are also identified. In field exercises, trainers and/or evaluators use the T&EOs to either identify areas of training needs or to more formally evaluate performance. When the T&EOs are used in an evaluation context, actions covered by individual items are evaluated by evaluators as satisfactory or unsatisfactory. A global assessment is then made of mission performance. In a training environment, the T&EOs can serve as a tool to aid the trainer in making diagnostic decisions concerning remedial training to remove performance weaknesses.

1-3. BACKGROUND OF ENGAGEMENT SIMULATION

ARI and private sector contractors have, under TRADOC sponsorship, been developing an approach to collective training which is called engagement simulation. Engagement simulation is a training concept guided by the assumption that a realistic simulation of the battle or or engagement provides best conditions for learning.

Engagement simulation fields opposing forces with conflicting missions in field exercises that incorporate free-play on the part of one or both of the forces. Judgments of tactical proficiency are made in terms of ground taken/held, casualties inflicted/suffered, and/or delay caused or experienced. The realistic simulation of the tactical environment is achieved through weapons *signature cues*, similar to what would be encountered in actual combat, through near real-time objective casualty/damage assessment and through free-play, where the engagement outcome is a direct function of opposing forces during the exercise.

In its evolution, engagement simulation has taken three forms: SCOPES, REALTRAIN, and MILES. The method originated with SCOPES (Squad Combat Operations Exercise, Simulated), which was an early version for infantry squad training. When expanded for combined arms tactical collective training, the approach became known as REALTRAIN. In both SCOPES and REALTRAIN, damage/casualty assessment is performed *manually* by a system of *maneuver controllers*. REALTRAIN has been successfully applied to cavalry and combined arms units of platoon and company (minus) size. With the introduction of laser equipment, damage/casualty assessment became more *automatic*, and the approach became known as MILES, an acronym for Multiple Integrated Laser Engagement System. As the engagement simulation method was evolved, it has been adapted for the training of larger units. It is believed that MILES will permit training of units as large as battalions.

Engagement simulation training attempts to elicit the same tactical behaviors as would occur in combat. The cues to which soldiers respond in the training approximate, as nearly as possible, those to which they would respond in battle. Combat elements attempt to respond in a manner appropriate to counter and neutralize the effects of the action of their adversaries. As tactical events occur, continual adjustments are made in the tactical plans which reflect the interplay of the opposing forces, and which represent adaptations to losses of men, material, and terrain as a result of enemy actions. Leaders and troops attempt to adapt their tactics to the battlefield situation of the moment.

CHAPTER 2
EVALUATION PROCESSES

2-1. THE BASIC EVALUATION PROCESS

Evaluation of performance is an integral and essential part of unit training and maintenance of readiness. Evaluation, as defined in the unit training context, can encompass activities best described as assessments that are made of a unit while undergoing training. Such assessments are internally generated and conducted. Evaluation can also pertain to external evaluations, sponsored by the next higher echelon.

Current Army practice uses both types of evaluations; both are needed. In the recent past, the external evaluation has been a vehicle for accountability.[1,2] Current FORSCOM guidance to active components (FORSCOM Regulation 350-1, 9 January 1979) states that though external (and internal) evaluations can contribute to the Unit Status Report, the primary aim of the external evaluation is the diagnosis of the unit's training proficiency early in a commander's tour in order to facilitate the use of the feedback in developing a training program. Under this guidance, the commanders are not required to assign ratings of satisfactory/unsatisfactory upon completion of external (or internal) evaluations or report results to Headquarters, FORSCOM.

Where the external evaluations may occur infrequently (e.g., once every 18 months), the internal evaluation is viewed by FORSCOM (Regulation 350-1, 9 January 1979) and by TRADOC (TRADOC Regulation 310-2, Draft, May 1978) as a continuous process by the trainers within the unit. It represents the day-to-day diagnosis and remediation of unit deficiencies.

---

[1]*Improved Army Training and Evaluation Program (ARTEP) Methods for Unit Evaluation—Volume I, Executive Summary: Study Design and Field Research.* ARI Technical Report TR-78-A26. November 1978.

[2]*Improved Army Training and Evaluation Program (ARTEP) Methods for Unit Evaluation—Volume II, Analysis.* ARI Technical Report TR-78-A-26. November 1978.

Regardless of the sponsor—internal or external—the process of assessment should be very similar. In this report we will treat them as equivalent.

The traditional approach to unit evaluation is guided by mission T&EOs; engagement simulation provides a different approach. Regardless of the approach, certain steps are always involved in the evaluation process. In any evaluation, there is a performing unit and a set of evaluators. During the evaluation process, the activities of the evaluated unit and those of evaluators are interwoven as follows:

a.   **Scenario Construction.** Prior to conduct of an evaluation, training managers specify a mission or missions, select ground, designate and construct a situation which will bring the unit in training against an opposing force in a simulated battle, and provide (or be prepared to provide) orders for the leader of the unit to be evaluated. The ground, the OPFOR, the unit mission(s), and orders—to coordination—make up the scenario. The scenario is designed to provide units an opportunity to exercise combat-relevant tasks.

b.   **Tasks Described by Doctrine.** There are tasks the unit is to perform in order to accomplish its mission, and guiding concepts as to how these tasks should be performed. Tactical doctrine provides the guidance as to what concepts are appropriate, and as to how they should be implemented. These concepts are translated into criteria. Criteria may refer to *products* of an activity or task to the *processes* that units should follow in order to attain desired goals. Criteria are used by evaluators and/or participants to evaluate performance.

c.   **Classification of Tasks.** Within the defined mission scenarios monitored by evaluators, tasks are performed by members of tactical units. These tasks may be classified in a variety of ways:

(1)   Mission phase chronology. Unit missions may be subdivided into a planning and an execution phase; they may be further broken out into convenient segments or phases that occ... :uring performance of a mission.

(2)   Task Content. The content of tasks to be performed may be classified in a variety of ways:

8

(a) Holistic or specific. The holistic perspective views performance of the unit as a whole—from without. A specific or instrumental perspective would be concerned with actions of specific unit elements.

(b) Internal coordination and support *versus* enemy referenced. Another system of classification would subdivide tasks into tasks of communication, tasks of support, and enemy referenced tasks. Tasks of communications involve the issue and dissemination of orders, the reporting of intelligence and tasks of lateral coordination. Support tasks involve providing logistical support, casualty evacuation, replacement of casualties, etc. A third set of tasks draws attention to actions taken in contact with an enemy force—essentially, the interactions between antagonists. Families of enemy-referenced tasks include:

- Finding the enemy force and estimating its character, size and disposition.

- Engaging the enemy force with weapons.

- Avoiding enemy observation/fire.

- Maneuver to include execution of overwatch.

d. **Observation of Performance**. Three next steps involve observation of performance, maintenance of a record or trace of performance, and evaluating performance observed. These steps are separable in concept. Depending, however, on the method of conducting exercises, these steps may be conducted separately or combined. We consider the first two together, then the evaluation of performance.

e. **Observation/Recording**: The performance of the unit—i.e., the *process*—or the *product* of that performance, or both are observed. A trace of what occurred—a history—is recorded either by evaluators or equipment. It is desirable that this record be preserved for further reference; otherwise, evaluation is entirely dependent on human memory. Not only is human memory fallible, but later events tend to be remembered better, so that significant events that occurred earlier are forgotten or distorted. Nonetheless, the memory of evaluators is always important. As discussed later, the memory of players or trainees can serve as valuable aids as well.

9

f.   **Evaluation**. The performance that was observed is compared with performance as it should have been as specified by doctrine; or, if product criteria are invoked, the product of the foregoing process is compared with the desired product or mission orders. In either case, doctrine provides, in effect, the templates that guide comparisons. These comparisons indicate the discrepancies between performance as observed, and doctrinally-approved performance. If the discrepancy is null, performance was effective. Typically, there are certain discrepancies, and these are detailed. They may be identified both within the time progression of events within the mission, and with regard to contents categories mentioned above.

g.   **Feedback and Use of Results**. Feedback is the communication of observed differences between actions observed and actions as they should have occurred. Feedback may be delivered either verbally or documented. It can be used as a short-term assessment within a training program or training exercise. It can be addressed to exercise participants, to include leaders, to trainers or to training managers. It can be specific or general. It can provide overall assessments of proficiency or of progress in long-term training programs. Feedback can contribute to both short-term and long-term planning of training.

2-2. ES AND CONVENTIONAL COMPARED

The six points described above outline a general framework for evaluation of tactical units. This frame can provide a means for comparison of engagement simulation with conventional methods of unit evaluation. By comparing the two methods in terms of this frame, differences between conventional and ES methods of evaluation become apparent. Identification of these differences helps define basic issues that must be resolved in integration of ES into ARTEP.

a.   **Scenario Construction and Exercise Structure**.

(1) Engagement simulation. In ES, scenarios are typically specified for both forces in a paired-missions context; meeting engagements, attack-defense, etc. However, scenarios serve only as points of departure for the rest of the exercise. Exercise outcome is not predetermined. Once the battle is joined, activities of units are dependent on actions and counter-

10

dependent upon the actual performance of the opposing forces in critical behaviors that emerg as a result of free-play between sides.

(2) T&E Outline format. In conventional exercises, the scenario is used as a stimulus situation for the performing unit. It is also a control framework for the exercise. Sequences of tactical events and OPFOR behavior are specific, controlled and linked closely to the sequences portrayed in the mission being evaluated. Missions, tasks and conditions of their performance are in effect predetermined. Within this controlled environment, there appears to be a limited set of behaviors which can be exhibited by the performing unit. In this sense the outcome of the exercise is predetermined.

b.   **Task Content and Performance**

(1)   Engagement simulation. Critical mission tasks and the conditions under which they are accomplished emerge from the free-play of the opposing forces. Taking and inflicting casualties are critical behaviors. Casualties serve as anchor points to identify and examine critical activities/behaviors of units that led to casualties taken or inflicted. Thus, critical behaviors are not predetermined but depend on the progress of the battle. Further, replays of the same scenarios, missions, and performing units can produce different critical behaviors for each exercise. Participants are typically well-motivated and eager.

(2)   T&E Outline format. Critical mission tasks, the conditions under which they are accomplished and standards defining acceptable performance for each task are preestablished and ordered. The exercise is structured to insure that these tasks are performed. One key evaluator/controller duty is to see that the stimulus—orders, OPFOR play—was presented as planned.

c.   **Observation and Control of Performance**

(1)   Engagement simulation. It is worth noting that engagement simulation uses *controllers* rather than observer/evaluators. The duties of controllers vary considerably as between REALTRAIN and MILES. In REALTRAIN, precedence is given to controllers

11

fastening the link as rapidly as possible between firers and targets. This is done by confirming that the target is in the firers sights and once the round is fired, communicating this information to the controller on the target vehicle. (This requirement is handled almost automatically in **MILES**.) A controller may be encouraged also to make notes as well about the positions and activities of firers/targets which proceeded the encounter. With either **REALTRAIN** or **MILES**, information as to firers/victims is relayed to a net control station which logs casualties and times. Basically, the controller is *not* regarded as an evaluator. His comments on circumstances relative to his element inflicting/suffering casualties may be solicited before the After-Action Review (**AAR**) to complete the record. The assumption explicit and implicit in the ES method is that casualties are the basic performance criteria, and that the key role of controllers is to assure that they are validly declared and sufficiently well documented.

(2) T&E Outline format. T&EO observers are assigned to units as evaluators and controllers. Their observations begin when the scenario OPORD is delivered by the senior controller. Throughout the exercise, they observe, record, and rate unit actions as they happen in the specified T&EO sequence. As controllers, they intervene where necessary to make the scenario happen as planned. Interventions are accomplished by issuing orders to the evaluated unit and by adjusting the movements and actions of the dedicated **OPFOR**. Casualties may be assessed to control the performing unit, or to be used as inputs to evaluate support elements concerned with casualty evacuation.

d. **Evaluation of Performance**

(1) Engagement simulation. ES performance evaluation measures have an objective base of casualties taken or inflicted as a result of interactions between elements of the two sides locked in battle. Casualties taken or inflicted are to serve as markers and as criteria for evaluation of prior tactical behaviors. If casualties are taken, there is an inherent implication that tactical behavior was faulty. It is further assumed that the performance of the men/ crew that inflicted the casualty was correct. Unlike the conventional method of evaluation performance assessment occurs not by an observer comparing performance with a preestablished doctrinal template, but through the action and counteractions between sides. These assessments by acts—kill or be killed—may be elaborated on in discussions in the AAR.

12

(2) T&E Outline format. In conventional exercises, T&E Outlines supply the mission-task framework for performance evaluation. Process behaviors and unit activities used in the outlines serve as cues to evaluators. Evaluators must supply the criteria from published tactics, doctrine and operational procedures in order to render judgments. The judgments are stochastically based. The evaluator must decide if, over many repetitions, the performance observed would contribute to or detract from successful mission performance. Thus, performance evaluation in the conventional sense requires subjective application of doctrinal criteria and subjective inferences about probable mission outcome based on the unit performance observed.

e.   **Data Reduction/Interpretation**

(1) Engagement simulation. To prepare for the After-Action Review (AAR), ES data reduction procedures are primarily integrative. At a Controller Debriefing, controller data is used to confirm data in the NCS log. Quickly, events in the NCS log are used to chronologically reconstruct the battle. Controllers add pertinent observations. The senior controller uses these comments and the NCS log as the AAR agenda.

(2) T&E Outline format. Determination of performance scores is done during observation, when satisfactory or unsatisfactory ratings are made. Summative ratings are made on tasks where more than one controller has observed unit performance or where a task has been repeated and rated several times. Although explicit instructions are not provided, presumably ratings of lower echelon units are summed to provide ratings for the next higher unit. Evaluators are required to infer from all task ratings if the mission was "successful." This is to be done in the absence of published guidance as to how to combine ratings.

f.   **Evaluation Feedback Delivery and Use**

(1) Engagement simulation. ES uses the AAR as its primary feedback vehicle. AARs stresses guided discussion of tactical events by participants. In this sense, the evaluation is accomplished by individuals examining their own performance and that of others. The process has been referred to in ARI documents as "discovery learning." Thus, the information is provided at the individual or squad/crew level. AARs results are designed for quick use in

successive training exercises. Post-exercise narratives have been used for observations of units over a series of exercises. Use of these narratives is said to provide excellent insights into what transpired during the exercise.

(2) **T&E Outline format.** Conventional exercises provide two forms of feedback, the critique and completed T&EO mission ratings. Critiques are conducted by evaluators and are directive. Comments are related to unit performance over all observed tasks. Individual performance is seldom mentioned. Critiques provide short-term information to participants and leader/trainers. T&E Outline ratings can provide information to trainers and training managers that is comparable across units and because of this, well adapted to longer term planning of training.

2-3. SUMMARY

Prior sections have developed steps required in any approach to collective training. They then focused on differences between engagement simulation as a methodology and typical practices in conventional (non-engagement simulation) exercises in the ARTEP program. It is important that these differences be recognized, for they help to define issues that must be resolved in the integration of the two approaches. In the next section, these differences are consolidated into *four* main issues that bear on conduct of unit proficiency assessments:

- Development of an integrated system of measures and criteria

- Allocation of functions for observation, judgment, and data collection.

- Development of data reduction/analysis procedures

- Development of procedures for feedback formulation and results utilization

These issues are quite similar to the issue areas reported on in phase one project work on conventional methods of unit training. The similarities between issues that bear on conducting effective conventional exercises, and melding conventional and ES exercises, pinpoint themes central to all unit training.

14

CHAPTER 3

ISSUES PERTAINING TO MEASUREMENT/CRITERIA

3-1. INTRODUCTION

An ES ARTEP involves integration of criteria from two approaches: the conventional method; and criterion measures as they have evolved from research on engagement simulation. Explored below are issues involved in assessment of criteria drawn from the two methods and in their integration into a single criterion system.

The criterion system should satisfy these requirements:

- Taken together, criteria should cover comprehensively critical behaviors, and products of these behaviors.

- Criterion measures should be valid.

- It must be possible to identify dimensions of behaviors subsumed under criteria.

- It should be possible to describe the *extent* to which behavior observed departed from that prescribed.

Satisfying all these conditions in unit evaluation is quite difficult. During long-term combat, many different behaviors may, and likely will, be important at one time or another. Many dimensions of tactical behavior are difficult to define and to quantify. Finally, objective evaluation of performance becomes more difficult if evaluations must depend on human observers.

These difficulties exist whether we refer to the conventional system of using ARTEP T&EOs or ES; but they arise in different degrees in these two criterion approaches. In melding the two approaches, we wish to look for strengths and weaknesses of each and how weaknesses of either approach may be complemented by the strengths of the other.

15

3-2. ADVANTAGES AND LIMITATIONS OF T&EOS AS CRITERIA

The ARTEP provides a framework of unit performance attributes by listing unit behaviors as tasks and subtasks for each mission in the order in which they are expected to occur. T&E items as criteria are the product of a task-analytic approach. They reflect the assumption that, if the behaviors as listed are satisfactorily performed (i.e. performed as specified by doctrine), on a stochastic basis, mission performance would be successful in combat. The T&E Outlines are presently used as guidance for conduct of unit training. As such, they offer a direct framework for assessment.

This frame is especially useful in providing guides for evaluation of all those activities that typically precede the battle—accepting orders, reconnaissance, formulating mission plans, and issuing orders.

Once the battle is joined, T&E Outlines for tactical units, as presently constructed and used, have several drawbacks as measurement devices. First, the processes listed under "Standards" are simply cues to trainers and evaluators. In many instances, all relevant dimensions of tasks are not defined, nor is guidance provided as to how tasks or attributes should be evaluated. Performance standards and, in come cases, task dimensions must be gleaned from field manuals. For example, one element in a T&E Outline for movement to contact is "use of bounding overwatch". Descriptions of correct use of bounding overwatch are not listed, but must be remembered from published guidance. Standards must be subjectively interpreted and applied by individual evaluators. Finally, the structure of T&E Outlines present mission performance steps in a straight chronological sequence. During conduct of the battle, it is unlikely that any given predetermined sequence of actions will be precisely appropriate. Rather, unit leaders and men must draw from their repertoire, responses appropriate to the demands of the situation at the moment. T&E Outlines as configured do not readily permit this flexibility.

16

Still other problems with T&EOs in their present form are evident, although these are in format and, in some cases, in lack of adequate instructions, rather than in concept.

- The present format of T&E items—evidently to save space—often combines two or more behavioral parameters under one item. In many items, these parameters can be orthogonal or independent. The evaluator can only mark SAT or UNSAT. Space allocated for marginal notes is minimal. A training manager, given completed T&E items, would not be able to "recapture" performance for each parameter. This means for formatting items is inimical to incisive training diagnosis.

- Often, evaluators are asked to make global judgements of unit performance without adequate guidance as to how to make their judgements.

- In some cases, instructions are insufficient; for example, items covering the use of cover/concealment by platoons, companies, and battalion. For a battalion field exercise, is the evaluation of use of cover/concealment a simple summation of use of cover/concealment by organic elements? If so, this should be stated in the instructions.

The points above should be regarded as technical problems rather than as faults of concept. All can be remedied. Nonetheless, until this is done, the full potential of T&EOs cannot be exploited as a training instrument.

3-3. ADVANTAGES AND LIMITS OF ES CRITERIA

If controllers are well trained and if their equipment (or the MILES system) is working, a major strength of engagement simulation is its objectivity in determining casualty and damage outcomes from conflicts. Further, when engagement outcomes are linked to unit and individual behaviors in post-exercise critiques, its power as a self— and peer—assessment tool is compelling.

Engagement simulation demands special equipment (scopes, and radios, or lasers/sensors) in order to more realistically play the battle. As such, it provides an environment in which, because of interactions between sides, unanticipated events can and

17

frequently do occur, and leaders and men must respond to the situation at the moment in intelligent and creative ways. They cannot fall back on prescribed T&E Outlines for the right answer. Its essential criteria are histories of casualties suffered by each side through time, which it produces, and the position of the unit(s) on conclusion of the exercise. In addition, unit positions through time can be recorded, as can traffic on tactical nets; but such records might be made—and might well be desirable—in conventional exercises as well. Thus, the unique character of engagement simulations is predicated on its ability to bring about casualties in a timely manner as a resultant of interactions between antagonists. (The fact that casualties are sometimes played in conventional exercises does not mean that the two approaches are comparable in casualty play; realistically, conventional exercises lack the means to play casualties well.)

a. **Product and Process**. In ES, casualties as *product* criteria provide the basis for conduct of the critique, or After Action Review (AAR), as it is called in engagement simulation. When the battle is over, the "killed" are resurrected, and the history of casualties is reviewed in platoon level exercises with all participants present. Each casualty becomes a "marker" which calls attention to the presumed effective behavior by the killer and ineffective behavior on the part of his victim. The AAR brings out these process behaviors or process *criteria as they bear on* actions of participants specific to each mini-engagement. Further, players and controllers participate. Information from all these sources are combined in group discussions to get the record straight. Thus, as actions emerge during the conflict, process criteria for evaluation of these actions are drawn out during the AAR. Since criteria represent the integration of inputs from players and controllers, they are apt to be more relevant and more valid than the rulings of a single evaluator using pre-formatted T&E Outlines.

b. **Casualties as Criteria**. Engagement simulation as a training method provides a far more realistic play of the battle than does the standard format. Casualties are produced by the interactions between sides. However, questions may be raised with regard to the sufficiency and (invariant) validity of the casualties as criteria. These questions apply in somewhat different form to front-line elements and to leaders.

18

(1) Casualty Criteria as Applied to Vehicles/Infantry. The most compelling criterion for battle success is kills or casualty ratios between sides. While casualty ratios have face validity in the gross sense, the occurrence of individual casualties cannot be taken, ipso facto, as proof that the casualty made a mistake. For example, in an attack mission, vehicles/infantry may make best use of *available* cover and concealment but still suffer casualties. Assuming that, in most instances, the player or the vehicle that became a casualty made a mistake, it is not *necessarily* true that he (it) did. Further, as casualties are not proof of ineffective performance, those who made avoidable errors are not always penalized by becoming casualties. Men of the opposing force may not be alert, or they may not be able to see soldiers/vehicles that needlessly exposed themselves. In sum, casualties are not an absolute indicator of effective or of ineffective performance of front-line elements.

(2) Casualty Criteria as Applied to Evaluation of Leader Performance. What, then, shall we say about use of casualties as indicators of leader performance? This question becomes increasingly pertinent as we go to company and battalion level ES exercises. Here, casualties may have occurred because of a poor plan, because of ineffective execution of the plan, or for both reasons. What observations of what elements are needed so that faults in planning can be separated from faulty execution? Another question: What is the role of casualties in evaluation of performance of companies and battalions? Here, with many men and vehicles engaged, AARs cannot well review casualties on a one-by-one basis. How can casualties be integrated and used as measures of leader/company performance? These are questions that remain to be clearly answered.

3-4. INTEGRATION OF CRITERIA: ES AND CONVENTIONAL METHODS

ES and conventional methods of evaluation introduce philosophically and operationally different approaches to problems of performance measurement. Questions arise as to whether the two can be successfully integrated, and, if so, on what basis.

19

The ES method requires added equipment. It may be justified on two bases:

- It provides a much more realistic better play of the battle.

- A National Training Center is being planned. Primary emphasis is on training of maneuver units as organic battalions and perhaps companies. The center will play two-sided field exercises using the ES format. The use of ES for home station training could serve better to prepare personnel for training at the NTC.

When the costs of added equipment (either scopes/radios or lasers, sensors) and the needs for skills to operate and maintain this equipment are considered, ES *cannot be used merely as an adjunct* to the conventional method that uses T&EOs as a guide.[3] Rather, conventional methods must be *built around* the ES frame and format. It further follows that, to the extent possible, training diagnosis and critiques must be built around, and support, casualty records. This raises at least two issues:

- How *complete* a coverage of critical actions does a valid and complete record of casualties provide? What supplementary information is needed for a complete training diagnosis? We can assume that casualties will be reported to an NCS. If a T&EO type format is used to permit collection of *additional* criterion information, what should its format be? How specifically should it try to describe behavioral parameters?

- What concepts shall be applied to collect additional criterion data? We can assume that casualties will be recorded at NCS. Once the supplementary information required (see above issue) has been determined, it can better be determined what added criteria are needed. One possibility, surely, is to use some format analogous to T&E Outlines to permit controllers to make evaluations. Here, since activities are apt to deviate both in their specifics and in time from those described in any prearranged outline, some flexibility will be needed. Possibly, as an aid to evaluation, a more general set of process

---

[3] This became obvious in observations of one battalion field exercise. Vehicles were numbered (as in REALTRAIN), but there was no net control station. Casualty records were than not available for the post-exercise critique.

20

criteria may be provided to evaluators. A broad classification of tasks such as that outlined under 2-1-c. *Classification of Tasks*, might be appropriate. Other possibilities include recording traffic on tactical nets—at least on a sampling basis—and maintaining a running record of positions of units and their elements over time.[4]

---

[4] The feasibility of relating such records to criteria is being pursued in ongoing ARI research.

CHAPTER 4

ISSUES COVERING OBSERVATION/CONTROL SYSTEM FUNCTIONS

Criterion measures for the integrated system determine largely what must be done to observe and record performance. The measures chosen specify what information is to be collected and roles of observers/evaluators in collecting this information and/or in helping to collect data.

However, other functions need to be performed in addition to those strictly related to the observation and evaluation of performance. These include control, rule enforcement, safety overwatch, administration and hardware system operation. Cumulatively, both types of functions define requirements placed on the observation/control in terms of numbers of people, equipment, skill and training, duties, coordination, and logistical support requirements.

Based on field observations, functions to be performed by controller/evaluators are as follows:

a. **Functions Required to Make ES Work.** There are certain things controllers must do just to make ES work. These functions differ between REALTRAIN and MILES. In REALTRAIN, controllers coordinate to apply the rules of engagement and declare casualties. This requires time and attention exactly when critical actions are taking place. It needs to be determined if ES controllers can simultaneously observe, rate, and record process-type behavior while accomplishing their ES functions. The result will have implications for both size and structure of the controller force.

b. **Functions Required for Performance Observation.** These activities bear on the assignment, location, and movements of observers. Coordination of activities between observers to keep track of coordination between elements of performing units that are not in physical contact are of interest. So are activities such as use of cover, concealment, and selection of routes for advance. Requirements in observing these

23

functions have implications for time sharing by observers between functions, as well as the number of observers needed.

c. **Functions Required for Performance Evaluation.** A distinction is made between this category and observation. Depending on the measures chosen, some data may be collected and analyzed later to determine proficiency. If this is the case, less skill and tr ing is needed for observers/data collectors. If each observer must also make judgments about what he sees, then considerably more training and skill will be required. This has implications for the size, structure, skill and training requirements of the observer group. Several things need to be determined:

- To what extent can information on certain measures be collected by individuals with no special training or qualifications?

- How much can observer training compensate for lack of prerequisite MOS or skills?

- What measures *must* be evaluated as they occur? Which ones can wait?

d. **Functions Required for Control.** As previously mentioned, ES emphasizes two-sided free play scenarios and allows situations to emerge. In certain instances, this approach may conflict with the need to structure the scenario to bring out preestablished assessment requirements. Here, depending on evaluation objectives, the scenario, or the action of OPFOR, or both may require more structure.[5] This creates the requirement for control-type functions not often used in ES. It must be determined how additional control functions fit with the other duties of exercise controllers and what increases in controller load they cause. To the extent possible, procedures would need to be developed so that personnel assigned as controllers could assume these additional duties without becoming severely overloaded.

---

[5] This is not without precedent even in straight ES "training" situations. Controllers may require that elements request permission to fire, if there are certain activities that the controller waits to observe or provide for the performing unit.

CHAPTER 5

ISSUES OF DATA REDUCTION,
ANALYSIS AND INTEGRATION

Procedures required to reduce, analyze and integrate data obviously depend on the nature of the data to be collected. Thus, issues such as those described above need to be resolved first. Lacking resolutions of these, we may start with certain "givens." ES will provide casualty records through time, and positions of opposing units, if desired. If the conventional method is used to supplement ES, we can expect evaluator ratings and/or observations. Other information such as records of tactical radio traffic may be collected as well. In any evaluation, then, we can safely assume a substantial amount of data will be collected. Further, this amount and its complexity will increase with unit size.

The information from these several sources needs to be consolidated. Further, if it is to be of most value to units in training, so they can review details of the engagement while it is fresh on their minds, integration must occur rapidly. Finally, speed of data reduction and integration must be balanced with the needs for feedback specificity and comprehensiveness. A number of questions then arise:

- How shall the data from the several sources be combined? By what logic?

- Can evaluator ratings be effectively used to better explain reasons for casualties and key events that preceeded them?

- Can the same formula or algorithms for data integration be used whatever the mission and in whatever way it unfolded?

- Are the data/information needed for the AAR and by the training manager the same, or do they differ? If so, in what respects?

These are among the issues that need resolution.

25

CHAPTER 6

ISSUES OF FEEDBACK AND RESULTS USE

6-1. INTRODUCTION

Post-exercise feedback may be provided to players and trainers as immediate knowledge of results, to training managers to help plan and schedule future training, and to headquarters personnel who may be concerned with strengths and weaknesses of institutionalized training programs.

Major differences between the engagement simulation approach to unit training and the conventional approach continue to be evident as one considers development of perscriptive guidance for feedback.

6-2. THE AFTER-ACTION REVIEW OR CRITIQUE

Thus far, ES has stressed quick verbal feedback essentially *between players* of the two sides with an AAR leader acting as "chairman of the board" and using casualty histories as agenda items. An exercise narrative is sometimes prepared for trainers and training managers; however, this feature does not appear to be unique to engagement simulation. Such a narrative could readily be prepared from well-kept T&EO protocols. The conventional exercise lends itself to a critique format, typically with a leader pointing out effective and ineffective behavior to participants.

6-3. USE OF AAR

As discussed elsewhere, the AAR appears to be a much more powerful teaching device than the typical critique given after conventional exercises. Casualties are confronted with consequences of their behavior rather than rulings of evaluators supported by tactical considerations that may well be stochastic and abstract. A major issue, however, should be considered before accepting the ES format fully. Since outcome data are so convincing, the potential exists for outcomes to overshadow correct and incorrect performances that are not

27

directly related to casualties as outcomes. Consider the following example. If a unit fails to use a sound tactical plan, but, by chance, surprises and defeats its adversaries, it becomes more difficult to illustrate performance deficiencies. It is important, then, when using ES techniques and AARs in unit assessment, that procedures be developed to put engagement outcomes in a proper perspective.

A further issue, raised above, is worth note. Namely, the time required to *integrate* data obtained from the two approaches may be such that delays of several hours would be encountered in assembly of information which is to be used in the AAR. Such delays in providing feedback should be avoided if at all possible in planning AARs for troops; delays are believed to be less harmful in planning AARs for leaders.

6-4. INFORMATION TO TRAINING MANAGERS

The ES method provides information to training managers as histories of kills and an exercise narrative; the conventional T&E format provides information in the form of completed checklists. Would training managers prefer one form of such information over another? Would they like both? This needs to be determined.

A further question may be raised; namely: Would training managers ever, for any purpose, wish to *compare* scores or quantitative data so as to *compare* performance of squads, platoons or companies? Since ES pits one unit against another, scores are *relative* to the capabilities of the two units. The T&EO format provides scores based on a common set of parameters. Thus, if the SATS/UNSATS from completed rating forms can be regarded as of some validity, then this format permits cross-unit comparisons. The ES format only permits comparisons of scores for the two units engaged.

28